Synthetic On-line Handwriting Generation by Distortions and Analogy

Harold MOUCHÈRE[1]*, Sabri BAYOUDH[2], Eric ANQUETIL[1] and Laurent MICLET[2]
[1]*IRISA, INSA de Rennes,*
[2]*IRISA, ENSSAT,*
Campus Universitaire de Beaulieu, Avenue du Général Leclerc, 35042 Rennes, France
harold.mouchere@irisa.fr; sabri.bayoudh@irisa.fr

Abstract. One of the difficulties to improve on the fly writer-dependent handwriting recognition systems is the lack of data available at the beginning of the adapting phase. In this paper we explore three possible strategies to generate synthetic handwriting characters from few samples of a writer. We explore in this paper both classical image distortions and two original ways to generate on-line handwritten characters: distortions based on specificities of the on-line handwriting and a generation based on analogical proportion. The experimentations show that these three approaches generate different distortions which are complementary. Indeed the combination of them allows the achievement, using only 4 original characters for the learning phase, of a mean of 91.3% of recognition rate for 12 writers.

1. Introduction

With the increasing use of the pen-based human computer interface, handwriting recognition systems have to become more and more accurate. One way to improve it is to adapt the system to the handwriting style of the current user as done in a previous work (Mouchère et al., 2007). The difficulty is to learn a new handwriting style from few samples.

In this paper we study three possible strategies to generate synthetic handwriting characters. The first one uses classical distortions from the off-line field. The two next ones are from our original contribution: the second one is based on the particularity of on-line handwriting and the third one is based on the use of analogical proportion on handwriting. Contrary to Varga et al. (2005) who generate data to learn a writer-independent system, the synthetic handwriting has to respect the writer script. The quality of this generation is evaluated by considering the performance of a writer-dependent simple classifier learned with synthetic characters generated from very few examples of his handwriting style.

We first present in the section 2 the classical image distortions. The section 3 focuses on the on-line handwriting distortions. The section 4 gives definition and properties of analogical proportion and shows how to use it to generate synthetic characters. The section 5 presents experimental results.

2. Synthesis by Image Distortions: Scaling and Slanting

Using synthetic data to learn a recognition system is mainly used with off-line systems which recognize an image of the character. The generation is done by using image distortion processes. For off-line character recognition, Cano et al. (2002) use distortions: slanting, shrinking, ink erosion and ink dilatation. Simard et al. (2003) extend the learning data base using elastic distortions on images to train a neural network. Moreover text lines can be distorted as in Varga and Bunke (2003) to train a Hidden Markov Model.

In this paper we limit us to scaling and slanting deformations because of the nature of our data. Indeed we deal with isolated on-line characters inputted directly by the user through a pen-based interface. Thus ink transformations, rotation, perspective and line distortions are not interesting in this case. Two random parameters correspond to the scale transformations, α_x and α_y which are the ratio of the corresponding scales. The slant allows the generation of inclined handwriting. It depends of one random parameter α_s which represents the tangent of the slant. A positive value slants the writing to the right and a negative to the left. The Figure 2 shows examples of distorted characters by scaling and slanting.

3. Synthesis by On-line Distortions

There are few works about using handwriting generation in order to increase on-line learning data. In the on-line field, the handwriting is considered as a parametric function $p_{(t)}$ corresponding to the pen positions. Varga et al.

*This work is supported by the Brittany Region and the CNRS.

(2005) use works about handwriting generation (Plamondon & Guerfali, 1998) to increase the size of a learning database to train an off-line sentence writer-independent recognizer. The authors use a unique on-line handwriting model which is distorted and joined to generate images of complete sentences. This approach is closed to our on-line distortion but unusable in our context because we need writer-dependent models. That is why we propose two simple distortions of on-line handwriting: *Speed Variation* and *Curvature Modification*

The aim of *Speed Variation* distortion is to modify the size of vertical and horizontal parts of the stroke, as shown by Figure 2, depending of a random parameter α_v. Indeed these straight parts of the writing can vary without changing the handwriting style. For this we modify the speed $\vec{V}_{(t)} = \big(x_{(t+1)} - x_{(t)}, y_{(t+1)} - y_{(t)}\big)$ depending of its direction. If this vector is near one of the axes then it is increased or decreased by the ratio α_v. The new synthetic handwriting is defined by $p'_{(t)}$:

$$p'_{(t)} = p'_{(t-1)} + \beta * \vec{V}_{(t-1)}, \text{ with } \beta = \begin{cases} 1 & \text{if } \arg(\vec{V}_{(t-1)}) \left[\frac{\pi}{2}\right] \in [\frac{\pi}{8}, \frac{3\pi}{8}], \\ \alpha_v & \text{else.} \end{cases} \quad (1)$$

The *Curvature Modification* distortion modifies the curvature of the writing as shown by Figure 2. It allows closing or opening the loops of handwriting. The curvature modification uses a random parameter α_c. The curvature at the point $p_{(t-1)}$ is defined by the angle $\hat{\theta}_{(t-1)} = (p_{(t-2)}, \widehat{p_{(t-1)}, p_{(t)}})$ in $]-\pi, \pi]$. In order to keep the structure of the character, we do not modify the straight lines and cusps. The equation 2 gives the position of the point $p'_{(t)}$ depending of the two previous points and of the original curvature $\hat{\theta}_{(t-1)}$ modified by α_c. The maximum angular modification is for $\hat{\theta}_{(t-1)} = \frac{\pi}{2}$.

$$(p'_{(t-2)}, \widehat{p'_{(t-1)}, p'_{(t)}}) = \hat{\theta}_{(t-1)} - \alpha_c * 4 * \frac{\left|\hat{\theta}_{(t-1)}\right|}{\pi} * (1 - \frac{\left|\hat{\theta}_{(t-1)}\right|}{\pi}). \quad (2)$$

4. Sequence Generation by Analogical Proportion

Analogy is a way of reasoning which has been studied throughout the history of philosophy and has been widely used in Artificial Intelligence and Linguistics (Lepage, 1998). We are interested here in a special case of analogy: the Analogical Proportion (AP) between four objects a, b, c and d in the same universe. Having four objects in AP is usually expressed as follows: "`a` *is to* `b` *as* `c` *is to* `d`". Depending on what are the objects, AP can have very different interpretations.When one of the objects is unknown, as in "`wolf` *is to* `leaf` *as* `wolves` *is to* `x`", finding a satisfying `x` is called *solving an analogical equation*. `x` would be `leaves` if we consider either a semantic interpretation of analogy as well as if the interpretation is on the combinatorics of the sequences of letters.

More formally, an *Analogical Proportion* (AP) on a set X is a subset of X^4. An element of AP writes $a : b :: c : d$, and reads: "`a` *is to* `b` *as* `c` *is to* `d`". As defined by Lepage (1998), an AP must verify:

Symmetry of the "as" relation:	$a : b :: c : d$	$\Rightarrow$	$c : d :: a : b$
Exchange of the means:	$a : b :: c : d$	$\Rightarrow$	$a : c :: b : d$
Determinism:	$a : a :: b : x$	$\Rightarrow$	$x = b$

4.1. Analogical Dissimilarity Between Four Objects

When four objects in X are not in AP, it may be interesting to introduce a quantity to measure how far they are from being in AP. This measure, called Analogical Dissimilarity (AD), has been proposed in Bayoudh et al. (2007). It has been constructed to verify properties coherent with the underlying AP. For example, if the set X is $\mathbb{R}^m$, we can define an AP by $a : b :: c : d \Leftrightarrow a + d = b + c$ and a coherent AD is such that:

1. $\forall u, v, w, x \in \mathbb{R}^m, AD(u, v, w, x) = 0 \Leftrightarrow u : v :: w : x$
2. $\forall u, v, w, x \in \mathbb{R}^m, AD(u, v, w, x) = AD(w, x, u, v) = AD(v, u, x, w)$
3. $\forall u, v, w, x, z, t \in \mathbb{R}^m, AD(u, v, z, t) \leq AD(u, v, w, x) + AD(w, x, z, t)$
4. In general, $\forall u, v, w, x \in \mathbb{R}^m : AD(u, v, w, x) \neq AD(v, u, w, x)$

Therefore, when defining $AD(a, b, c, d) = \delta(a + d, b + c)$, where δ is a distance in $\mathbb{R}^m$, it is easy to prove that the four properties above still hold true.

4.2. Analogical Dissimilarity Between Four Sequences

Let us define a sequence as an ordered set of objects, these objects being for example vectors in $\mathbb{R}^m$, or more generally elements of an alphabet Σ. The set of sequences on Σ is denoted $\Sigma^\star$.

$$
\begin{array}{ll}
u = & 9 \backsim 9 \backsim 9\,9\,9\,9\,9 \backsim E\ 1\ 2 \backsim 4\ L\ 6\ 9\ 9\ 9 \\
v = & 1\ L \backsim 8\ 9\,9\,9\,9\,9\,10\ E \backsim 2\ 2\ 4\ L \backsim 8\ 8\ 9 \\
w = & \backsim \backsim 9\ \ 8\ 9\,9\,9\,9\,9\,10\ E\ 2\ 2\ 3\ 3\ L\ 8\ 9\ 9 \backsim \\
x_0 = & 1\ L \backsim 8\ 9\,9\,9\,9\,9\,10\ E\ 2\ 2\ 3\ 3\ L\ 8\ 8\ 8 \backsim
\end{array}
$$

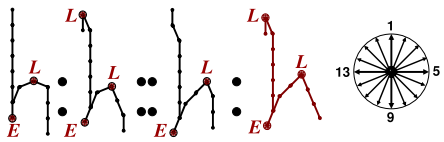

Figure 1. Resolution on Freeman direction sequences by AP and the corresponding characters.

We assume that there exists an analogical dissimilarity AD on Σ. To extend to sequences, we augment Σ to Σ' by adding a special symbol $\backsim$ and we consequently augment the AP and the definition of AD (Delhay & Miclet, 2004). We define an *alignment* between four sequences of $\Sigma^\star$ as the result of inserting some $\backsim$ symbols in the four sentences to give them the same length on Σ'. Then, the cost of an alignment between these four sequences is the sum of the analogical dissimilarities between the 4-tuples of letters given by the alignment.

We basically represent characters by Freeman code sequences. Let $\Sigma' = \{1, 2, .., 16, \backsim\}$ be the augmented Freeman code alphabet, with an AP composed of equations like: $3 : 5 :: 12 : 14$. Figure 1 presents four sequences corresponding to the pictures on the right side. These sequences are composed of elements of Σ' plus capital letters representing anchorage points that will be described in the next section.

We define the Analogical Dissimilarity between four Sequences $ADS(u, v, w, x)$ in $\Sigma^\star$, as the cost of an alignment of minimal cost of the four sequences. We have given an algorithm to compute the ADS between four sentences and another to compute the k sequences x that give the lowest value to $ADS(u, v, w, x)$, given u, v and w. We have shown that ADS has the same properties as the AD between objects except for the third property (section 4.1 point 3). The sequence x_0 in Figure 1 is the solution to the equation $\underset{x}{\boldsymbol{Min}}\ ADS(u, v, w, x)$.

4.3. Sequence Segmentation

Aligning sequences composed only of Freeman code could lead to match strokes of different types and therefore to a non pertinent generation of the fourth stroke. In order to generate a coherent sequence we choose to use anchorage points. Practically, the set of Freeman codes and the set of anchorage points belong to the same space, where the intra-distance between the Freeman codes is lower than the intra-distance between the anchorage points, which is lower than the inter-distance between the sets. Hence, resolution by AP will first resolve analogies on anchorage points then resolve analogies on sequences between anchorage points as shown in the Figure 1. Capital letters in sequences represent different anchorage points (like E for cups, L for y-extremum), those points are represented by a bold point. The anchorage points are chosen in an appropriate way to conduct a logical modeling of letters with respect to the most stable strokes of each letter class (Anquetil & Lorette, 1997).

5. Experimentations

Twelve different writers have made 40 times the 26 lowercase letters (1040 characters) inputted on a PDA. Each writer database is randomly split in four parts with 10 characters per class. We use them in a 4-fold stratified cross validation: one fourth for $D10$ database (260 data) and three fourth for $D30$ database (780 data). The experimentations are made of two phases in which simple writer-dependent systems based on Fuzzy Inference Systems (Mouchère et al., 2007) are learned.

Firstly two writer-dependent classifiers are learned for each writer on his $D10$ database and evaluated on his $D30$ database, and inversely. The two mean recognition rates are the Reference Rates $RR10$ and $RR30$, i.e. the recognition rate achievable without character synthesis. Here the $RR10$ is 82.3 % and $RR30$ is 94.5 %.

Secondly the handwriting generation strategies are tested. For a given writer, two, three, four, six, eight or ten characters per class are randomly chosen in his $D10$ database. Then 300 synthetic characters are generated per class to make a synthetic learning database. A writer-dependent classifier is learned with this base and tested on the $D30$ database of the writer. This experiment is done 3 times per cross-fold. The mean and deviation of these 3×4 performance rates are computed. Finally the means of these measurements are computed and correspond to the Writer Dependent Mean recognition rate (WDM) and the Writer Dependent standard Deviation (WDD).

We study four different strategies for the generation of synthetic learning databases. The strategy *"Image Distortions"* chooses randomly for each generation one among the three image distortions. By the same way the strategy *"On-line & Image Distortions"* chooses randomly one distortion among the image distortions and on-line distortions. The *"Analogy and Distortions"* strategy generates two thirds of the base with the previous strategy and the one third with analogical proportion. Table 1 sums up the experiment results.

Nb. of used characters	2	3	4	6	8	10
Image Distortions	76.1 ± 3.3	82.5 ±2.4	85.8 ±2.0	89.4 ±1.7	91.5 ±1.6	92.7 ±1.3
On-line & Image Distortions	84.4 ±2.6	88.0 ±2.1	90.3 ±1.7	92.3 ±1.6	93.4 ±1.2	94.2 ±1.0
Analogy and Distortions	84.9 ±2.6	89.3 ±2.1	91.3 ±1.4	93.5 ±1.1	94.5 ±1.0	95.2 ±0.9

Table 1. Writer Dependent Mean (WDM) performance rate (%) and Writer Dependent Deviation (WDD) (%) for different synthetic handwriting generation strategies, the reference rate $RR10$ is 82.3 % $RR30$ is 94.5 %.

Two main conclusions can be deduced from these results. Firstly the aim of the article is achieved. Indeed with only two original characters the "Analogy and Distortions" allows better writer-dependent recognition rate than the reference $RR10$ and eight to be better than $RR30$. Secondly Table 1 shows that the three generation approaches (image distortions, on-line distortions and analogy) are complementary. Indeed the three strategies generate 300 data and we note that the richer are the distortions the better are WDM rates. Furthermore, Figure 2 shows that each strategy allows different variants of original characters not achievable by other ones.

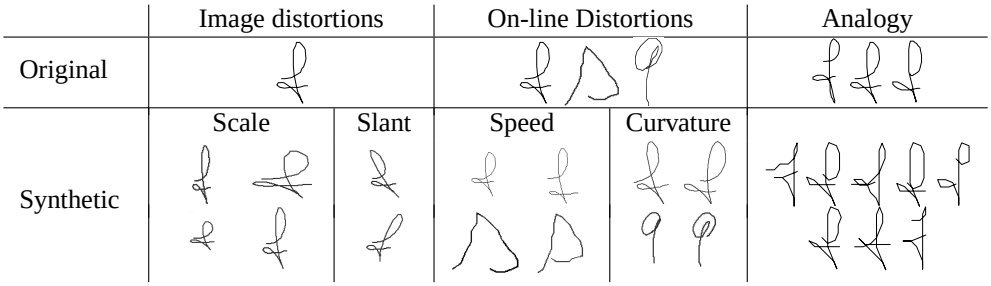

Figure 2. Examples of synthetic characters generated by the three approaches.

6. Conclusion

We have shown that our two original generation strategies bring new distortions with regards to the classical image modifications. Furthermore, the combination of these three strategies keeps the writer scripts as it allows to learn efficient writer-depend recognizer with synthetic data generated from very few original characters.

Future works can focus on two goals: to use the generation of handwriting with delta-lognormal synergies (Plamondon & Guerfali, 1998); to improve the quality of synthetic data generating by analogy to avoid some very distorted characters which can impair the recognition.

7. References

Anquetil, E., & Lorette, G. (1997). Perceptual model of handwriting drawing application to the handwriting segmentation problem. In *Proc. of the 4th int. conf. on document analysis and recognition* (p. 112-117).

Bayoudh, S., Miclet, L., & Delhay, A. (2007). Learning by analogy : a classification rule for binary and nominal data. In *Proc. of the int. joint conf. on artificial intelligence* (Vol. 20, pp. 678–683).

Cano, J., Prez-Cortes, J.-C., Arlandis, J., & Llobet, R. (2002). Training set expansion in handwritten character recognition. In *Proc. of the 9th int. workshop on structural and syntactic pattern recognition* (p. 548-556).

Delhay, A., & Miclet, L. (2004). Analogical equations in sequences : Definition and resolution. In *Proc. of international colloquium on grammatical inference* (pp. 127–138).

Lepage, Y. (1998). Solving analogies on words: an algorithm. In *Proc. of coling-acl'98* (Vol. 1, pp. 728–735).

Mouchère, H., Anquetil, E., & Ragot, N. (2007). Writer style adaptation in on-line handwriting recognizers by a fuzzy mechanism approach : The adapt method. *Int. Journal of Pattern Recognition and Artificial Intelligence, 21*(1), 99-116.

Plamondon, R., & Guerfali, W. (1998). The generation of handwriting with delta-lognormal synergies. *Biological Cybernetics, 78*, 119-132.

Simard, P., Steinkraus, D., & Platt, J. (2003). Best practice for convolutional neural network applied to visual analysis. In *Proc. of the 7th int. conf. on document analysis and recognition.*

Varga, T., & Bunke, H. (2003). Generation of synthetic data for an hmm-based handwriting recognition system. In *Proc. of 7th int. conf. on document analysis and recognition* (p. 618-622).

Varga, T., Kilchhofer, D., & Bunke, H. (2005). Template-based synthetic handwriting generation for the training of recognition systems. In *Proc. of 12th conf. of the international graphonomics society* (p. 206-211).